

Lensing Degeneracies Revisited

Prasenjit Saha

Astronomy Unit, School of Mathematical Sciences
Queen Mary and Westfield College
London E1 4NS, UK

This paper shows that the mass-sheet degeneracy and other degeneracies in lensing have simple geometrical interpretations: they are mostly rescalings of the arrival-time surface. Different degeneracies appear in Local Group lensing and in cosmological lensing, because in the former the absolute magnification is measured but the image structure is not resolved, whereas in the latter the reverse usually applies. The most dangerous of these is a combination we may call the ‘mass-disk degeneracy’ in multiply-imaging galaxy lenses, which may lead to large systematic uncertainties in estimates of cosmological parameters from these systems.

Subject headings: gravitational lensing

To appear in AJ, Oct 2000

1. Introduction

A curious feature of gravitational lensing is that most of the observables are dimensionless. This fact leads to some scaleabilities in lensing theory, which show up as parameter degeneracies when interpreting observations. These degeneracies were analyzed in detail in Gorenstein et al. (1988, hereafter G88), elaborating on Falco et al. (1985). Since that time, while lensing theory has not changed much the observational situation has changed greatly—recall that in 1988 cluster lensing was still controversial, and Milky Way microlensing was several years in the future—and with it the emphasis of theory has shifted. So it is interesting at this time to rederive the G88 degeneracies, discuss their current observational context, and try to gain some new insights into the old results.

G88 discussed three basic degeneracies, which they called the similarity, prismatic and magnification transformations, and combinations of them. The most subtle of these is the magnification transformation; it seems to have been independently discovered at least two more times, and is now usually called the mass-sheet degeneracy. The present paper will be about the same transformations, but unlike G88 who started from the lens equation, we will think about transformations of the arrival-time surface.

2. The degeneracies

In fact, lensing degeneracies can all be interpreted as simple transformations of arrival-time surface

$$t(\boldsymbol{\theta}) = t_{\text{geom}} + t_{\text{grav}}. \quad (1)$$

and we can recognize three kinds.

- ‘Similarity transformations’ scale both t_{geom} and t_{grav} by a constant factor. Such operations scale time delays between images but leave image positions and magnifications unchanged.
- The ‘mass-sheet degeneracy’ mixes t_{geom} and t_{grav} but in such a way that $t(\boldsymbol{\theta})$ is scaled by a constant factor. This multiplies time delays and all magnifications by a constant factor but leaves image positions and *relative* magnifications unchanged.
- Various other transformations can be written down that modify t_{grav} and possibly also t_{geom} in various ways, but leave $t(\boldsymbol{\theta})$ and its derivatives unchanged at all image positions. These will have no effect on observables, but they cannot be ignored because they imply uncertainties in what can be inferred about lenses from the observables.

To derive these degeneracies, we start with the full expression for the arrival time

$$t(\boldsymbol{\theta}) = \frac{1}{2}(1 + z_L) \frac{D_L D_S}{c D_{LS}} (\boldsymbol{\theta} - \boldsymbol{\beta})^2 - (1 + z_L) \frac{8\pi G}{c^3} \nabla^{-2} \Sigma(\boldsymbol{\theta}), \quad (2)$$

where $\boldsymbol{\theta}$ and $\boldsymbol{\beta}$ are angular positions on the image and source planes respectively, ∇^{-2} denotes the inverse of a two-dimensional Laplacian, and the other symbols have their usual meanings.

2.1. The similarity transformations

The simplest of the degeneracies appears if the distance factor is Eq. (2) is unknown (through uncertainty in one or more of D_S , D_L or cosmology), which allows the transformation

$$\frac{D_L D_S}{D_{LS}} \rightarrow s \frac{D_L D_S}{D_{LS}}, \quad \Sigma(\boldsymbol{\theta}) \rightarrow s \Sigma(\boldsymbol{\theta}). \quad (3)$$

(Here and below s is an arbitrary constant.) G88 call (3) a similarity transformation. The only effect on observables is to multiply time delays between images by s ; neither image positions nor magnifications change.

If the images are not resolved, another similarity transformation,

$$\boldsymbol{\theta} \rightarrow \sqrt{s} \boldsymbol{\theta}, \quad \boldsymbol{\beta} \rightarrow \sqrt{s} \boldsymbol{\beta}, \quad \Sigma(\boldsymbol{\theta}) \rightarrow s \Sigma(\boldsymbol{\theta}), \quad (4)$$

(not explicitly considered by G88) becomes possible. Here both sources and images are rescaled by \sqrt{s} , so magnifications are unaffected, while once again time delays get multiplied by s . To avoid a degeneracy of names, I suggest calling Eq. (3) a ‘distance degeneracy’ and Eq. (4) an ‘angular degeneracy’, reserving ‘similarity transformations’ for the whole category.

The distance and angular degeneracies are independent, in the sense that it is possible to break one without breaking the other. Clearly, one can combine this pair to invent other pairs of independent similarity transformations. One such pair, which I suggest calling the ‘parallax’ and ‘perspective’ degeneracies, are motivated as follows.

Consider the effect of parallax, i.e., moving the observer. Say the observer moves transverse to the optical axis by \mathbf{r}_{obs} . For the observer, the lens will move by $-\mathbf{r}_{\text{obs}}/D_L$ and the source by $-\mathbf{r}_{\text{obs}}/D_S$, which amounts to keeping the $\boldsymbol{\theta}$ fixed and moving $\boldsymbol{\beta}$ by $\mathbf{r}_{\text{obs}} D_{LS}/(D_L D_S)$. Applying this change to the arrival time (2) and discarding terms with no $\boldsymbol{\theta}$ -dependence gives

$$t(\boldsymbol{\theta}) = (1 + z_L) \left[\frac{D_L D_S}{c D_{LS}} \left(\frac{1}{2} \boldsymbol{\theta}^2 - \boldsymbol{\theta} \cdot \boldsymbol{\beta} \right) - \frac{1}{c} \mathbf{r}_{\text{obs}} \cdot \boldsymbol{\theta} - \frac{8\pi G}{c^3} \nabla^{-2} \Sigma(\boldsymbol{\theta}) \right]. \quad (5)$$

If \mathbf{r}_{obs} is known and non-zero, the transformations (3) and (4) are not allowed individually, but the mixture

$$\boldsymbol{\theta} \rightarrow s \boldsymbol{\theta}, \quad \boldsymbol{\beta} \rightarrow s \boldsymbol{\beta}, \quad \frac{D_L D_S}{D_{LS}} \rightarrow s^{-1} \frac{D_L D_S}{D_{LS}}, \quad \Sigma(\boldsymbol{\theta}) \rightarrow s \Sigma(\boldsymbol{\theta}) \quad (6)$$

is still possible. Here the magnifications will depend on \mathbf{r}_{obs} , but Eq. (6) does not change them because it rescales $\boldsymbol{\theta}$ and $\boldsymbol{\beta}$ equally. I suggest calling Eq. (6) the perspective degeneracy because it preserves the product of the distance and angular scales. Meanwhile, the similarity transformation

$$\boldsymbol{\theta} \rightarrow s \boldsymbol{\theta}, \quad \boldsymbol{\beta} \rightarrow s \boldsymbol{\beta}, \quad \frac{D_L D_S}{D_{LS}} \rightarrow s \frac{D_L D_S}{D_{LS}}, \quad \Sigma(\boldsymbol{\theta}) \rightarrow s^3 \Sigma(\boldsymbol{\theta}) \quad (7)$$

is independent of Eq. (6) and we can think of it as the degeneracy that is *broken* by a parallax observation, so I suggest calling it the parallax degeneracy.

One usually factors out the similarity transformation by working with a scaled arrival time surface like so

$$\tau(\boldsymbol{\theta}) = \frac{1}{2} (\boldsymbol{\theta} - \boldsymbol{\beta})^2 - 2 \nabla^{-2} \kappa(\boldsymbol{\theta}). \quad (8)$$

Here the scaled arrival time τ , the scaled surface density (or convergence) κ and the operator ∇^{-2} are all dimensionless. The physical arrival time and density are

$$t(\boldsymbol{\theta}) = (1 + z_L) \frac{D_L D_S}{c D_{LS}} \tau(\boldsymbol{\theta}), \quad \Sigma(\boldsymbol{\theta}) = \frac{c^2}{4\pi G} \frac{D_S}{D_{LS} D_L} \kappa(\boldsymbol{\theta}). \quad (9)$$

The usual lensing potential is $\psi = 2 \nabla^{-2} \kappa$ and the bending angle is $\boldsymbol{\alpha} = \nabla \psi$.

2.2. The mass-sheet degeneracy

We now rewrite (8) by discarding a $\frac{1}{2}\beta^2$ term, since it is constant over the arrival-time surface, and using $\nabla^2\theta^2 = 4$, to get

$$\tau(\theta) = 2\nabla^{-2}(1 - \kappa) - \theta \cdot \beta. \quad (10)$$

The transformation

$$1 - \kappa \rightarrow s(1 - \kappa), \quad \beta \rightarrow s\beta. \quad (11)$$

clearly just rescales time delays while keeping the image structure the same; but since the source plane is rescaled by s all magnifications are scaled by $1/s$, leaving relative magnifications unchanged. The effect on the lens is to rescale the lensing mass and then add or subtract a constant mass sheet. G88 call (11) a magnification transformation, but ‘mass-sheet degeneracy’ is its usual name nowadays.

For a circular lens, the mass-sheet degeneracy preserves the total mass inside an Einstein radius θ_E . We can see this by invoking the two-dimensional analog of Gauss’s flux law in electrostatics, which in lens notation becomes

$$\oint \alpha \times d\mathbf{l} = 2 \int \kappa d^2\theta, \quad (12)$$

or that the normal component of α , integrated along any closed loop, is proportional to the enclosed mass. Along an Einstein ring, α is always radial and hence normal to the ring; also, its magnitude always equals θ_E (since a source at the centre is imaged onto the ring). Hence, the left hand integral in Eq. (12) depends only on θ_E . Meanwhile the right hand integral gives twice the enclosed mass. Thus, fixing the Einstein radius fixes the enclosed mass.

The mass-sheet degeneracy is broken if there are sources at more than one redshift. The reason is that we can no longer factor out the source-redshift dependence as we did in Eqs. (8) and (9). Instead, we can replace (8) and (9) with

$$\tau(\theta) = \frac{1}{2}(\theta - \beta)^2 - 2\frac{D_{LS}}{D_S}\nabla^{-2}\kappa(\theta), \quad t(\theta) = (1 + z_L)\frac{D_L}{c}\tau(\theta), \quad \Sigma(\theta) = \frac{c^2}{4\pi G}\frac{1}{D_L}\kappa(\theta), \quad (13)$$

and replace Eq. (10) with

$$\tau(\theta) = 2\nabla^{-2}\left(1 - \frac{D_{LS}}{D_S}\kappa\right) - \theta \cdot \beta, \quad (14)$$

Sources at different redshifts imply simultaneous equations of the type (14) but with different factors of D_{LS}/D_S , which prevents a transformation like (11).

2.3. Other degeneracies

G88 discuss one other transformation, which they call ‘prismatic’, consisting of adding the same constant to both the source position and the bending angle. Physically, this amounts to adding a very massive lens at very large transverse distance while pushing the source in the opposite direction. So it is not as important as the similarity and magnification transformations.

Clearly, one can concoct any number of localized transformations that leave $t(\theta)$ and its derivatives unchanged at all image positions and do not make κ negative anywhere. An obvious one is what we may call a ‘monopole’ transformation: any circularly symmetric redistribution of mass inwards of all observed images, and any circularly symmetric change in mass outside all observed images will have no effect on observables. A more subtle example, which causes an ambiguity between close and wide binary lenses in Local Group lensing, is discussed by Dominik (1999).

The monopole transformation has an important indirect effect: it changes the mass-sheet degeneracy into a ‘mass-disk degeneracy’—as long as the disk is larger than the region of images, a circular disk and an infinite sheet are equivalent in lensing—and a much more dangerous effect, since it cannot be eliminated by the requirement that κ goes to 0 at large θ . Figure 1 illustrates.

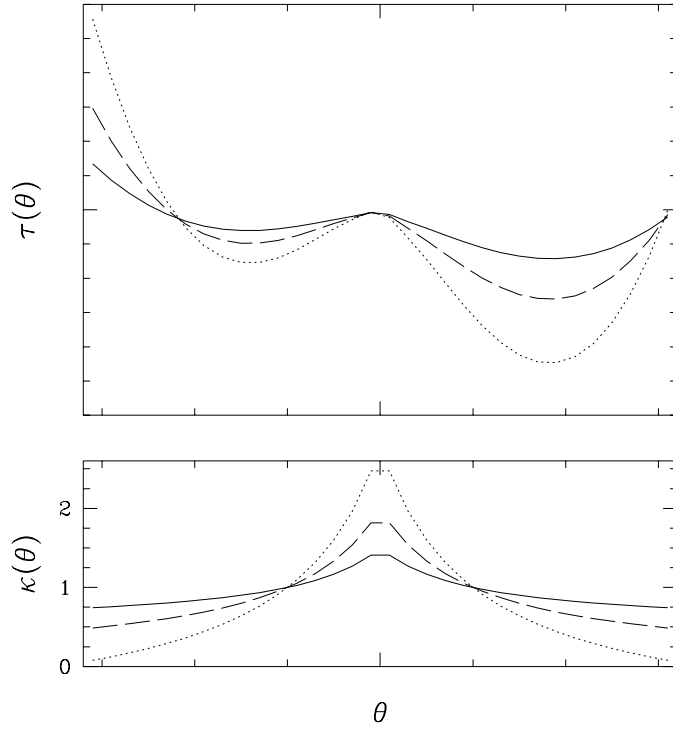


Figure 1. Illustration of the mass-disk degeneracy, showing the surface density (lower panel) and the arrival time (upper panel) for three circular lenses. The units, except for κ , are arbitrary. The arrival time indicates a saddle point (looking like a local minimum in this cut), a maximum, and a minimum. The dashed curves correspond to a non-singular isothermal lens. Stretching the time scale amounts to making lens profile steeper (dotted curves) and shrinking the time scale amounts to making the lens profile shallower (solid curves). Note that there is a limit to stretching, because otherwise κ will become negative somewhere in the region with images—negative κ *outside* that region can always be avoided by adding an external monopole. But there is no limit to shrinking.

3. Digression: velocity dispersions

Though not a lensing observable, velocity dispersion is often measured in connection with lensing, and is worth discussing here.

In a lens having approximately critical density, a typical internal velocity v satisfies

$$v \sim c\theta_E^{\frac{1}{2}}. \quad (15)$$

To derive this, we write R for the lens's size and M for its mass, and recall that $\theta_E \sim R/D_L$ for critical density, $v^2 \sim GM/R$ from the virial theorem, and that $\theta_E \sim GM/(c^2 D_L)$.

The most familiar example of the scaling (15) is for an isothermal lens. As is well known, this lens derives from a stellar dynamical sphere with constant velocity dispersion σ : the density is $\rho = \sigma^2/(2\pi Gr^2)$, which amounts to a projected density of $\Sigma = \sigma^2/(2GD_L^2\theta^2)$, leading to

$$\theta_E = 4\pi \frac{\sigma^2}{c^2} \frac{D_L}{D_S}. \quad (16)$$

One can use Eq. (16) to define a formal σ for any approximately circular lens. This formal σ can usefully serve as a surrogate for θ_E . Moreover, because of the relation (15) the formal σ will be of order the internal velocities in the lens, but in general it will be different from the actual velocity dispersion.

To elaborate, let us consider the relation between observed velocity dispersions and mass distribution. For a stellar system with no rotation or other streaming motions, the virial theorem states that $\langle v^2 \rangle = \langle \mathbf{r} \cdot \nabla \Phi \rangle$, where v is the stellar velocity, the averages $\langle \dots \rangle$ are over the stellar distribution function, and Φ is the total gravitational potential. Thus far there are no symmetry assumptions. If, however, spherical symmetry does apply then the line-of-sight direction must contribute the same as the orthogonal directions, and hence

$$\langle v_{\text{los}}^2 \rangle = \frac{1}{3} \langle \mathbf{r} \cdot \nabla \Phi \rangle, \quad (17)$$

v_{los} being the line-of-sight stellar velocity. If Φ is due to an isothermal sphere with dispersion σ then $\mathbf{r} \cdot \nabla \Phi = 2\sigma^2$ everywhere, which gives

$$\langle v_{\text{los}}^2 \rangle = \frac{2}{3} \sigma^2 \quad (18)$$

(cf. equation 4.6 in Kochanek 1993). For other spherical lenses one may compare $\langle v_{\text{los}}^2 \rangle$ with the formal σ^2 derived from Eq. (16). For example, consider a homogeneous sphere of stars of radius R , with no non-stellar matter. Using (17), we have

$$\langle v_{\text{los}}^2 \rangle = \frac{GM}{5R} \quad (19)$$

where M is the total mass. If this sphere is barely compact, M and R satisfy $4GM D_{\text{LS}} / (c^2 D_{\text{L}} D_{\text{S}}) = \theta_{\text{E}}^2$ and $R/D_{\text{L}} = \theta_{\text{E}}$. Inserting these values into (19) and then eliminating θ_{E} using (16) gives

$$\langle v_{\text{los}}^2 \rangle = \frac{\pi}{5} \sigma^2, \quad (20)$$

only 6% different from (18).

The above relations imply two things: (i) Eq. (17) indicates that there is a considerable range allowed in $\langle v_{\text{los}}^2 \rangle$ for lenses with given formal σ , and (ii) Eq. (20) shows that even if $\langle v_{\text{los}}^2 \rangle$ is observed to have the expected isothermal value, it does not follow that the lens is isothermal. And all this uncertainty is present without even considering ellipticity and velocity anisotropy.

In summary, for galaxy and cluster lenses an order-of-magnitude relation of the type

$$\theta_{\text{E}} \simeq 2'' \times \frac{\langle v_{\text{los}}^2 \rangle}{(300 \text{ km s}^{-1})^2} \quad (21)$$

is useful, but velocity dispersion is not a precise constraint unless the mass distribution is already known. Perhaps lenses become much better constrained if there is much more detailed velocity information; the answer seems unknown, but see Dejonghe & Merritt (1992) and Romanowsky & Kochanek (1999).

4. Appearances of the degeneracies in Local Group Lensing

In Local Group lensing, the similarity transformations are relevant. The mass-disk degeneracy does not apply because absolute magnifications are always measured, and anyway there are no disk-like lens components involved.

In most Local Group microlensing events only the magnification as a function of time is measured; the distances are unknown and the image structure is unresolved, so both distance and angular degeneracies (alternatively, both parallax and perspective degeneracies) apply. Note that the events being time-dependent and hence furnishing a whole sequence of arrival-time surfaces does not prevent the similarity transformations—for each event one can scale the whole sequence of arrival-time surfaces by the same factor.

In a few cases (~ 15 out of ~ 500 events observed so far) one degeneracy has been broken through additional observational information, and there are prospects for breaking the degeneracies completely in future with the help of observations from satellites. The requirements for degeneracy-breaking are well known, but it is interesting to interpret them in terms of the similarity transformations.

4.1. Proper motions

Proper motion measurements break the angular degeneracy (4), leaving the spatial degeneracy (3).

The significance of proper motion measurements was already appreciated by Refsdal (1966b), though the configuration then envisaged (lensing of visible one star by another, with separate proper-motion measurements for both) is not now considered realistic. A more realistic situation, independently pointed out by Gould (1994a), Nemiroff & Wickramasinghe (1994), and Witt & Mao (1994), is of a lens transiting the source star, in which case the finite size of the source will flatten the peak of the light curve; if the angular size of the source star can be estimated then $d\beta/dt$, and hence θ_E , can be inferred. Alcock et al. (1997) observed such an event. Another situation which enables $d\beta/dt$ to be measured is when not the lens itself but a caustic of a binary lens crosses the source. Albrow et al. (1999, 2000a,b) Afonso et al. (2000) and Alcock et al. (2000) have made such measurements.

4.2. Parallaxes

For a parallax observation, one needs to introduce the effect of a suitable \mathbf{r}_{obs} in Eq. (5), and this can be brought about in two ways. One way, suggested by Refsdal (1966b) and Gould (1992, 1994b, 1995), is to have separate observers, using one or more satellites. The other way, suggested by Gould (1992), is to exploit the Earth's acceleration. Now, a constant $d\mathbf{r}_{\text{obs}}/dt$ is irrelevant in Eq. (5) because it can be absorbed inside $d\beta/dt$. But a known $d^2\mathbf{r}_{\text{obs}}/dt^2$ modifies both the magnification (photometric parallax) and the proper motion of the image centroid (astrometric parallax). Photometric parallax events have been observed by Alcock et al. (1995), Bennett et al. (1997) and Mao (1999).

Parallax observations leave the perspective degeneracy (6), which holds the combination

$$\frac{D_L D_S}{D_{LS}} \Sigma(\boldsymbol{\theta}) \quad (22)$$

constant. For a single mass, (22) is $\propto \tilde{r}_E^2$ (\tilde{r}_E being the Einstein radius projected onto the observer plane). Thus we recover the well-known result that parallax measurements determine \tilde{r}_E .

4.3. Prospects for combining proper motion and parallax

As will be clear from the above (and elsewhere—Refsdal 1966b already made this point) combining proper motion and parallax measurements will lift all the degeneracies, and enable the lens mass to be solved for completely. Prospects for combined measurements have been discussed in several papers. Paczyński (1998) and Boden et al. (1998) suggest using interferometry to measure both proper motion and astrometric parallax, while Miyamoto & Yoshii (1995), Berlinski & Saha (1998) and Gould & Salim (1999) advocate using interferometry for proper motions and photometry for parallaxes.

5. Appearances of the Degeneracies in Cosmological Lensing

The degeneracies in cosmological lensing are complementary to those in Local Group lensing. The angular degeneracy does not appear because there is always some resolved image structure. The distance degeneracy appears, but in a very simple way—with redshifts usually measurable, the distance factor in Eq. (2) is $\propto H_0^{-1}$ times a weak and readily-quantifiable dependence on cosmology. The main thing to worry about is the mass-disk degeneracy. The following very briefly discusses the various contexts.

5.1. In quasar microlensing

In quasar microlensing the mass-disk degeneracy is actually useful! Modeling microlensing of lensed quasars involves computing lightcurves in a potential of stars plus smooth matter. In such computations (e.g., Refsdal & Stabell 1997) a standard trick uses the mass-disk degeneracy to transform away the effect of the smooth matter by rescaling the stellar masses appropriately—see Eq. (24) of Paczyński (1986), which appears to be an independent discovery of the degeneracy.

The angular degeneracy is also present, because while one obviously cannot change the angular scale of the macro-images, it is not forbidden to rescale the micro-image system within each macro-image, along with the source’s proper motion.

5.2. In cluster lensing

Another independent discovery of the the mass-disk degeneracy, this time in the context of cluster lensing, was by Schneider & Seitz (1994). Kaiser’s (1995) formula expressing $\nabla \ln(1 - \kappa)$ in terms of observable ellipticities makes the degeneracy particularly explicit: multiplying $(1 - \kappa)$ by a constant will not change the ellipticities. These and later papers have drawn considerable attention to the need to break the degeneracy, and research towards this end is active. Most of the effort is directed towards using the magnification information from number counts of background galaxies (see e.g., Taylor et al. 1998), but AbdelSalam et al. (1998) use the information that comes from having a range of source redshifts.

5.3. In quasar macrolensing

Although lensing degeneracies were originally discovered in the context of quasar macrolensing, recent literature in this area (e.g., the article on ‘Modeling Galaxy Lenses’ by Blandford et al., 2000) usually does not discuss degeneracies. The reason, perhaps, is that the popular parametrized lens models have focused attention on their respective parameters and away from the global transformations that produce degeneracies.

Considerable work has been done on fitting parametric models to the detailed image structure (e.g., Kochanek 1995, Kochanek et al., 2000). Such work often produces precisely constrained values for the radial density gradient and the core radius. But—and this is very important—those values are conditional upon particular parametrized lens models, because (a) the mass-disk degeneracy allows one to change the radial density gradient drastically without changing the image structure at all, and (b) the monopole degeneracy makes core radii if anything more free. Nor, as we saw in the previous section, do velocity dispersion measurements provide strong independent constraints unless the mass profile is assumed already known.

Thus, the mass-disk and monopole degeneracies point to some significant uncertainties in our current knowledge of galaxy lens profiles, and hence to uncertainties in estimates of cosmological parameters from quasar lensing.

The effect of the mass-disk degeneracy on estimates of h was already fully appreciated in Falco et al. (1985). Given a lens model that reproduces all observations of a lensed quasar and its host galaxy, one is still free to stretch or shrink the scale of the arrival-time surface (i.e., h^{-1}) using the mass-disk degeneracy—see Figure 1. There is a limit to stretching, because eventually κ somewhere will reach zero; this means that there is an upper limit on the inferred h . There is no limit to shrinking the arrival-time surface: the lens can get arbitrarily close to a disk with $\kappa = 1$ and the inferred h will get closer and closer to zero! To prevent this happening one must incorporate some assumptions about the steepness of the mass profile. Model-builders are familiar with such behavior (see e.g., Wambsganss & Paczyński 1994, Williams & Saha 2000).

Degeneracies are even more dangerous for inferences of Ω and Λ from lensing, because individual lenses contain no information on these parameters, only the ensemble of lenses does.¹ Several researchers (Maoz

¹ In a little known companion paper to the famous Refsdal (1964) on time delays and h , Refsdal (1966a) suggested that time-delays for systems at different redshifts could be put on a sort of Hubble diagram to determine the other cosmological parameters. But at present, researchers prefer to fit the redshift-dependencies of the density of multiple-image systems and the distribution of image separations, which also depend on cosmology; these two quantities are much easier to observe than time delays, but more awkward to interpret because magnification bias enters.

& Rix 1993, Kochanek 1996, Park & Gott 1997, Chiba & Yoshii 1999) have attempted to constrain Ω and Λ from the redshift-dependence of the source density or the image separations, or both. The results are conditional upon different assumptions made by the various authors, and in particular upon very specific lens profiles. Williams (1997) studies the dependence of the image-separation statistics on lens profiles, and concludes that it is much larger than the dependence on cosmology. For example, by making the lens profiles less steep in the inner regions she can make small-separation systems more magnified and hence (because of magnification bias) more abundant at high redshifts, thus completely drowning out the effect of cosmology.²

6. Summary

Lensing degeneracies can be simply understood as rescalings or other transformations of the arrival-time surface that leave various image properties unaffected. The most important of these are as follows.

- ‘Similarity transformations’ typically arise in microlensing. There are two independent ones which, depending on context, are usefully taken as:
 1. a ‘distance’ degeneracy where the distance scale varies while the angular scale stays fixed; and
 2. an ‘angular’ degeneracy where the angular scale varies while the distance scale stays fixed;
 or as
 1. a ‘perspective’ degeneracy where the product of the distance and angular scale varies while the ratio stays fixed; and
 2. a ‘parallax’ degeneracy where the ratio stays fixed while the product varies.
- The ‘mass-disk degeneracy’ is typical of cosmological lensing. It rescales

$$1 - \frac{\langle \text{density} \rangle}{\langle \text{critical density} \rangle}$$

within a finite disk larger than observed region, in the process rescaling the total magnification and the time delays, but otherwise leaving images unaffected.

- ‘Localized’ degeneracies do not change the arrival-time surface at image positions, but change it elsewhere; no lensing observable is altered, but other properties such as core radii and velocity dispersions may be.

The most insidious of these is the mass-disk degeneracy when it appears in multiply-imaging galaxy lenses, where it translates into a serious source of uncertainty in estimates of h from time delays, and even worse uncertainties in estimates of Ω and Λ from image statistics.

I am grateful to the referee for a number of detailed comments and suggestions.

² The image-separation statistics are complicated by the existence of a number of wide-separation quasar pairs at low redshifts, with no visible lens. These are currently thought to be binary quasars with no lensing involved (Kochanek et al. 1999), though spectral similarities in some cases cast doubts upon that interpretation (Small et al. 1997, Peng et al. 1999). Park & Gott (1997) find that the image-separation statistics with the wide-separation pairs included as lenses is not reproducible using power-law lenses. Williams (1997) finds that the same statistics can be reproduced if the lensing galaxies have changing logarithmic density profiles and follow the scaling laws characteristic of spirals rather than ellipticals, and concludes that the resolution of the nature of the wide-separation pairs as lenses or binaries will lead to constraints on the lensing population. Kochanek et al. (1999) say that Williams’s examples are “inconsistent with the known properties of galaxies and lenses” but do not explain which known properties.

References

- AbdelSalam, H.M., Saha, P., & Williams, L.L.R. 1998, AJ, 116, 1541
- Afonso, C. et al. 2000, ApJ, 532, 340
- Albrow, M. et al. 1999, ApJ 522, 1011
- Albrow, M. et al. 2000a, ApJ 534, 894
- Albrow, M. et al. 2000b, [astro-ph/0004243](#)
- Alcock, C., et al. 1995, Apj 454, L125
- Alcock, C., et al. 1997 ApJ 491, 436
- Alcock, C., et al. 2000 ApJ [astro-ph/9907369](#)
- Bennett, D.B. 1997, BAAS 191, 8303
- Berlinski, P. & Saha, P. 1998, New Ast Rev, 42, 111
- Blandford, R., Surpi, G., & Kundić, T. 2000, in Brainerd, T.G. & Kochanek, C.S. Eds., “Gravitational Lensing: Recent Progress and Future Goals”, [astro-ph/0001496](#)
- Boden, A.F., Shao, M., & Van Buren, D. 1998, ApJ 502, 538
- Chiba, M. & Yoshii, Y. 1999, ApJ, 510, 42
- Dejonghe, H., & Merritt, D. 1992, ApJ 391, 531
- Dominik, M. 1999, A&A, 349, 108
- Falco, E.E., Gorenstein, M.V., & Shapiro, I.I. 1985, ApJ 289, L1
- Gorenstein, M.V., Falco, E.E., & Shapiro, I.I. 1988, ApJ 327, 693
- Gould, A. 1992, ApJ 392, 442
- Gould, A. 1994a, ApJ 421, L71
- Gould, A. 1994b, ApJ 421, L75
- Gould, A. 1995, ApJ 441, L21
- Gould, A. & Salim, S. 1999, ApJ 524, 794
- Kaiser, N. 1995, ApJ 439, L1
- Kochanek, C.S. 1993, ApJ 419, 12
- Kochanek, C.S. 1995, ApJ 445, 559
- Kochanek, C.S. 1996, ApJ 466, 638
- Kochanek, C.S., Falco, E.E., & Muñoz, J.A. 1999, ApJ 510, 590
- Kochanek, C.S., Keeton, C.R., McLeod, B.A. 2000, [astro-ph/0006116](#)
- Mao, S. 1999, A&A, 350, L19
- Maoz, D., & Rix, H.-W. 1993, ApJ 416, 425
- Miyamoto, M., & Yoshii, Y. 1995, AJ 110, 1427
- Nemiroff, R.J. & Wickramasinghe, W.A.D.T. 1994, ApJ 424, L21
- Park, M.-G., & Gott, J.R. III 1997, ApJ 489, 476
- Peng, C.Y. et al. 1999, ApJ 524, 572
- Paczynski, B. 1986, ApJ 301, 503
- Paczynski, B. 1998, ApJ, 494, L23
- Refsdal, S. 1964, MNRAS 128, 307
- Refsdal, S. 1966a, MNRAS 132, 101
- Refsdal, S. 1966b, MNRAS 134, 315
- Refsdal, S. & Stabell, R. 1997, A&A, 325, 877
- Romanowsky, A.J., & Kochanek, C.S. 1999, ApJ 516, 18

- Schneider, P., & Seitz, C. 1995, A&A, 294, 411
- Small TA, Sargent WLW, Steidel CC 1997, AJ 114, 2254
- Taylor, A.N., Dye, S., Broadhurst, T.J., Benitez, N., & van Kampen, E. 1998, ApJ 501, 539
- Wambsganss, J., & Paczyński, B. 1994, AJ 108, 1156
- Williams, L.L.R. 1997, MNRAS 292, L27
- Williams, L.L.R. & Saha, P. 2000, AJ 119, 439
- Witt, H.J. & Mao, S. 1994, ApJ 429, 66